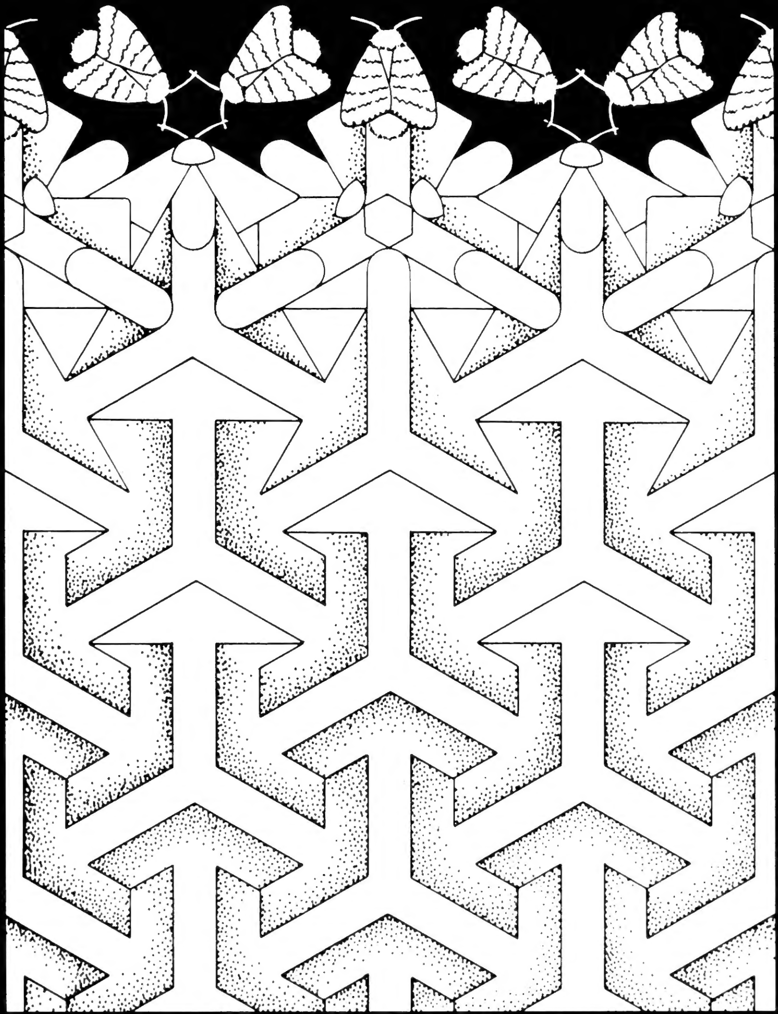
Predicting the Susceptibility of Illinois Forest Stands to Defoliation by the Gypsy Moth

David L. Swofford, Michael R. Jeffords, and Karen W. O'Hayer



Illinois Natural History Survey Biological Notes 131
April 1988

Illinois Natural History Survey, Lorin I. Nevling, Chief
A Division of the Illinois Department of Energy and Natural Resources

Printed by authority of the State of Illinois

No charge is made for most publications of the Illinois Natural History Survey, and a list of those publications is available upon request. Single copies of most publications are available to anyone requesting them. Requests for multiple copies should be made in writing and should explain the use to be made of the publications. Address correspondence to the Office of the Chief at the address below.

Illinois Natural History Survey
Natural Resources Building
607 East Peabody Drive
Champaign, Illinois 61820

65653–2M–4–88
US ISSN 0073-490X

Cover drawing: John Sherrod
Graphic assistance: Connie Leverich
Composition: Patty Duzan and Eva Steger
Editor: Audrey Hodgins

Citation:
Swofford, D.L., M.R. Jeffords, and K.W. O'Hayer. 1988. Predicting the susceptibility of Illinois forest stands to defoliation by the gypsy moth. Illinois Natural History Survey Biological Notes 131. 4 p.

Predicting the Susceptibility of Illinois Forest Stands to Defoliation by the Gypsy Moth

David L. Swofford, Michael R. Jeffords, and Karen W. O'Hayer

Abstract

A logistic regression equation that may prove capable of predicting the susceptibility of Illinois forest stands to defoliation by the gypsy moth is discussed and demonstrated. This equation incorporates easily measured variables for which values can be found in existing forest inventories. The equation is then applied to a set of sample data available through the Illinois Forest Inventory Data Processing System. Susceptibility predictions for a given area within the state can be obtained from the Illinois Natural Resources Information System at the Illinois Natural History Survey via modem and a local phone line (217-244-1361) or a toll-free number (1-800-247-4647).

Introduction

The gypsy moth (*Lymantria dispar*) has been a major forest pest in the eastern United States for more than a century. During that time, certain forest stands within its range have rarely been defoliated while others have suffered frequent and severe defoliation. Bess et al. (1947) emphasized the relationship between conditions at a given site and susceptibility or resistance to defoliation by gypsy moths. Among these conditions are soil type, litter depth, dominant species of trees and ground cover, growth characteristics of the timber, canopy cover, topography, and level of disturbance. The presence or absence of such structural features of trees as bark flaps and tree holes, which provide above-ground resting and pupation sites for gypsy moths, is also important (Campbell et al. 1975a, 1975b, 1977).

During the past decade, researchers have begun to develop models capable of predicting the susceptibility or resistance of forest stands to defoliation by gypsy moths. Houston and Valentine (1977), for example, relied on principal components analysis ordinations. They compared forests with known defoliation histories on the basis of the presence and number of certain structural features and found that stands with similar susceptibilities tended to cluster in ordinations. Later, they used discriminant analysis to classify stands as susceptible or resistant to defoliation (Valentine and Houston 1979). By collecting data on structural features, species, and tree size from stands of known or suspected susceptibility, they were able to compute variables that reflected the quality and quantity of gypsy moth habitat. The best subset of discriminatory variables was then used to compute a linear discriminant function by which each stand was classified as susceptible or resistant. Using this method, Valentine and Houston correctly classified 107 of 121 stands.

Because the variables used in the discriminant function were expensive to collect and difficult to measure, Valentine and Houston (1984) developed two other discriminant functions. The first relied on a simplified subset of the variables used in their earlier studies but still required data on structural features. The second used data on tree species and size only, information that is typically available from standard forest inventories. The first function, which included variables related to structural features, proved substantially superior to the second.

Gypsy moths were first detected in Illinois in 1973. Since then, small infestations have been found in a number of counties, mainly in the northeastern part of the state. These isolated populations have been reduced through mass trapping, the use of pesticides, and other control measures. About 3.9 million acres or 11 percent of the total land in Illinois is forest (Pelz and Thom 1977), most of which is managed for wildlife, recreation, or the production of timber. Because of the potential danger that the gypsy moth poses to Illinois forests, a quantitative method for identifying stands having the greatest risk of defoliation is desirable. In this paper we describe a risk assessment model that we believe makes the best use of the forest inventory data currently available for Illinois.

Development of the Model and Results

Because the gypsy moth is not yet widely established in Illinois, no intrastate data on which to base a model for predicting forest susceptibility are available. Consequently, we developed a model based on information from forest stands in the northeastern United States, stands that have a known history of susceptibility or resistance to defoliation by the gypsy moth. Data were provided by Dr. Harry Valentine of the U.S. Forest Service for 121 stands that had been classified as susceptible or resistant, primarily on the

Dr. David L. Swofford is an Assistant Biochemical Systematist with the Faunistic Surveys and Insect Identification Section of the Illinois Natural History Survey. Dr. Michael R. Jeffords, Assistant Entomologist, and Karen W. O'Hayer, Associate Research Biologist, are members of the Section of Economic Entomology at the Survey.

basis of their defoliation histories (Valentine and Houston 1979). Each tree in sample plots in these stands had been identified by species, measured for size, and assessed for condition.

Ideally, we would have used Valentine and Houston's (1984) preferred function to calculate for each stand a discriminant score that could then be used to assess the risk to Illinois forests. This function, however, included variables (sums of diameters of trees with bark flaps and with bark fissures) for which no information is presently available in data bases for Illinois stands, and the resampling required to obtain it was neither economically possible nor logistically feasible. We therefore formulated the best possible model for predicting the susceptibility of stands in the northeastern United States that relied only on those variables for which we have Illinois data. This model can now be applied directly to data from land surveys conducted by the Forestry Division of the Illinois Department of Conservation (IDOC). State-owned public lands, which include state parks, historic sites, and conservation and fish and wildlife areas, are surveyed once every ten years. Private lands are surveyed at the request of the owner or when they are to be included in a study conducted by the IDOC or by another agency of the state. The information gleaned from these surveys is tabulated in the Illinois Forest Inventory Data Processing (IFIDAP) system (Pelz and Thom 1977).

Variables for which data are available for stands in the northeastern United States and in Illinois include number of trees, basal area of each species (on a per acre basis), and average condition of trees (either by species or overall). (Some question remains, however, regarding the comparability of the relatively subjective "tree condition" ratings, within the IFIDAP system itself and between the IFIDAP and the Valentine and Houston samples.) Ratings of moisture (wet, moist, dry) and topography (flat, rolling, rough) were also assigned to each stand based on soil and topographic maps. Finally, tree species were classified according to their desirability as food for gypsy moth larvae (preferred, intermediate, nonpreferred). This classification followed Valentine and Houston (1979) wherever possible; however, tree species of unknown preference were assigned the classification of similar (usually congeneric) species.

A variety of techniques for discriminating between susceptible and resistant stands was evaluated, including linear classificatory discriminant analysis (e.g., Cooley and Lohnes 1971), logistic regression (Walker and Duncan 1967), and nearest-neighbor (nonparametric) discriminant analysis (Cover and Hart 1967). Although linear discriminant analysis is the most powerful of these methods when the assumptions of multivariate normality and equality of covariance matrices hold (Efron 1975), other methods may be justified when these assumptions are violated as severely as they were by our data sets. In particular, logistic regression is suitable for a variety of underlying assumptions (Anderson 1972), and its superiority for nonnormal distributions has been demonstrated empirically (Press and Wilson 1978).

For each method, variables were selected subjectively and via stepwise algorithms. The model that we found most suitable is the separate sample logistic regression equation (Anderson 1972):

$$\text{Probability (susceptibility)} = \frac{1}{1 + \exp(3.659 - 0.0216\,\mathrm{NPRF} + 0.0140\,\mathrm{BAPRF})}$$

where NPRF is the number of live trees (diameter at breast height ≥ 2.5 cm) per acre in the preferred food class, and BAPRF is the total basal area (ft^2) per acre of those trees. This function, plotted in Figure 1, primarily reflects the importance of the density of preferred trees. A tendency for the probability of susceptibility to decline as the ratio of basal area to the number of preferred trees increases is also evident, in accord with the findings of Valentine and Houston (1979, 1984). This tendency suggests that forests in early successional stages are more susceptible to defoliation by gypsy moths than are mature forests.

This simple logistic regression identified susceptibility among the 121 forest stands in the northeastern United States nearly as well as a number of more complicated models that we evaluated. If we consider a stand "susceptible" when its estimated (posterior) probability of susceptibility exceeds 0.5, 65 of 77 sus-

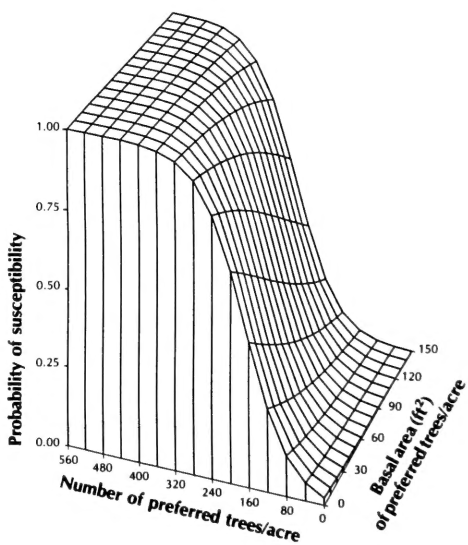


FIGURE 1. Plot of the logistic regression equation used to predict the susceptibility of forest stands to gypsy moth defoliation.

ceptible stands were correctly identified (15.6% misclassified) and 36 of 44 resistant stands were correctly identified (18.2% misclassified). By comparison, Valentine and Houston's (1984) preferred discriminant function yielded apparent misclassification rates of 5.2 percent and 6.8 percent for susceptible and resistant stands, respectively. This function, however, included bark-related variables that were not available in data bases for Illinois forests. When Valentine and Houston (1984) calculated a second function that used only standard inventory variables, apparent misclassification rates were 16.9 percent (susceptible) and 15.9 percent (resistant). Thus, our function performs about as well as did Valentine and Houston's second function.

The misclassification rates given above were computed by evaluating the performance of a function on the same data set used in its construction and, consequently, are termed "apparent." These rates are potentially misleading because they may not reliably estimate misclassification rates when the function is applied to new data sets. To estimate how well the selected function would predict the susceptibility of previously unclassified stands, we performed a cross-validation analysis (Stone 1974). Each of the 121 stands in the northeastern United States was deleted in turn, and the logistic regression function was fit using data from the remaining 120 stands. This function was then used to predict the susceptibility of the omitted stand. The results of the cross validation are encouraging because no additional misclassifications occurred.

The logistic regression equation was used to evaluate each of 176 stands in Illinois for which IFIDAP data were available. Of these, 14 were classified as susceptible. The low susceptibility of the stands in this sample may have occurred because most of them are mid-to-late successional forests. A representative sample of the results is shown in Table 1.

Discussion

We hasten to emphasize that the quantitative method described here does not take into account a number of aspects of gypsy moth biology that should be considered when evaluating the susceptibility of a given stand. Such physiographic variables as topography and moisture may be particularly important (Bess et al. 1947; Houston and Valentine 1977) and were not included in our final function because of our concern over subjective scoring. In any case, the improvement in the fit of the model provided by these additional variables was not judged substantial enough to justify their inclusion. In addition, the degree to which results from stands in the northeastern United States can be extrapolated to climatic conditions of the Midwest remains unknown.

We began this study with the intention of developing an objective function that could serve as a first step in predicting susceptibility of Illinois forest stands to defoliation by gypsy moths. We present our results in the hope that others engaged in similar endeavors will find them useful. The accuracy of our predictions cannot as yet be evaluated; however, as the gypsy moth continues to spread into Illinois, we will be able to improve our predictions concerning where the pest is most likely to become a serious problem.

Access to Prediction Information

Information generated by this Illinois gypsy moth model is available through the Illinois Natural Resources Information System (INRIS) located on the Prime computer of the Department of Energy and Natural Resources housed at the Illinois Natural History Survey in Champaign. INRIS grants users access to various data sets related to the natural resources of Illinois (McReynolds 1986) via terminals or personal computers. Within Illinois, INRIS data are available via modem and a local phone line (217-244-1361) or a toll-free line (1-800-247-4647).

TABLE 1. Estimated probability of susceptibility to defoliation by gypsy moths for selected Illinois forest stands. Stands with values >0.5 are considered susceptible.

	Number of preferred trees/acre	Basal area (ft^2) of preferred trees/acre	Estimated probability of susceptibility
Baldwin Lake and Softwood	12.430	18.750	0.025
Castle Rock Fearer Tract	155.556	48.182	0.275
Chain O'Lakes	98.657	58.529	0.088
Ferne Clyffe Stand 2	256.874	78.333	0.689
Fox Ridge State Park	87.087	44.828	0.083
Dolan Lake Area 2	147.392	42.000	0.257
Horseshoe Lake	32.608	46.087	0.027
Marshall County Conservation Area	125.894	39.286	0.184
Sand Prairie Scrub Oaks Pine Plantation	0.000	0.000	0.025
Sand Prairie Scrub Oaks Cruise	213.131	65.244	0.509
Trail of Tears State Forest Subcompartment 2	485.686	102.000	0.995
Piney Creek Nature Preserve	134.674	54.211	0.182
Oquawka State Wildlife Refuge	11.067	8.571	0.028
Railsplitter State Park	38.845	6.170	0.052
Mississippi Palisades	154.230	54.882	0.251

Output from the model is available by county or by susceptibility level. Figure 2 shows a representative county screen. In this example, the user has selected "Hardin County" from a menu of Illinois county names for which data exist. In this instance, a single screen accommodates the data; for other counties, more than one screen may be required and the user elects whether or not to continue the display. The total acreage of forest stands surveyed in a county appears near the bottom of the screen as does the total acreage of susceptible stands in that county. After displaying the data for a county, the user may select another county or return to INRIS.

The model will be updated and expanded as new forest inventories are generated by the Illinois Department of Conservation and the U.S. Forest Service. For more information on INRIS or on how to obtain predictions from the Illinois gypsy moth model, readers may contact INRIS Director Mark McReynolds by mail at 172 Natural Resources Building, 607 East Peabody Drive, Champaign, IL 61820 or by phone at 217-333-6006.

Acknowledgments

The authors thank Harry Valentine of the U.S. Forest Service for data on the forest stands of the northeastern United States and the Illinois Department of Conservation, Forestry Division, for making their forest inventory data available to us. Three staff members of the Illinois Natural History Survey also merit acknowledgment: Ed Kornkven for formatting the original data and submitting them to the model; Lane Smith for assisting with statistical calculations; and Mark McReynolds for providing information on INRIS. Finally, we thank David Onstad, William Ruesink, Peter Bayley, and Ronald Larkin of the Survey; Marc Linit, University of Missouri, Columbia; and Phillip Watson, Ferris State College, Big Rapids, Michigan, for helpful reviews of the manuscript. This research was partially supported by contract no. IP2-997 from the Illinois Department of Energy and Natural Resources.

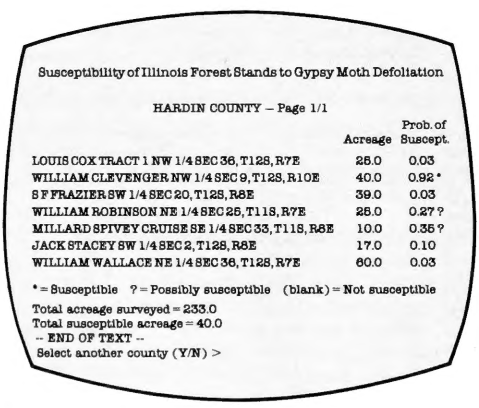


FIGURE 2. Sample INRIS screen: prediction of the susceptibility of forest stands in Hardin County, Illinois, to gypsy moth defoliation.

Literature Cited

Anderson, J.A. 1972. Separate sample logistic discrimination. Biometrika 59:19–35.

Bess, H.A., S.H. Spurr, and E.W. Littlefield. 1947. Forest site conditions and the gypsy moth. Harvard Forest Bulletin 22. 56 p.

Campbell, R.W., D.L. Hubbard, and R.J. Sloan. 1975a. Patterns of gypsy moth occurrence within a sparse and numerically stable population. Environmental Entomology 4:535–542.

Campbell, R.W., D.L. Hubbard, and R.J. Sloan. 1975b. Location of gypsy moth pupae and subsequent pupal survival in sparse, stable populations. Environmental Entomology 4:597–600.

Campbell, R.W., R.J. Sloan, and C.E. Biazak. 1977. Sources of mortality among late instar gypsy moth larvae in sparse populations. Environmental Entomology 6:865–871.

Cooley, W.W., and P.R. Lohnes. 1971. Multivariate data analysis. John Wiley and Sons, New York. 364 p.

Cover, T.M., and P.E. Hart. 1967. Nearest neighbor pattern classification. IEEE Transactions on Information Theory 13:21–27.

Efron, B. 1975. The efficiency of logistic regression compared to normal discriminant analysis. Journal of the American Statistical Association 70:892–898.

Houston, D.R., and H.T. Valentine. 1977. Comparing and predicting forest stand susceptibility to gypsy moth. Canadian Journal of Forestry Research 7:447–461.

McReynolds, M. 1986. The Illinois Natural Resources Information System (INRIS). Illinois Natural History Survey Reports 259:3–4.

Pelz, D.R., and R.H. Thom. 1977. The Illinois forest inventory data processing system. Illinois Department of Conservation Management Notes 2.

Press, S.J., and S. Wilson. 1978. Choosing between logistic regression and discriminant analysis. Journal of the American Statistical Association 73:699–705.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society, Series B, 36:111–147.

Valentine, H.T., and D.R. Houston. 1979. A discriminant function for identifying mixed-oak stand susceptibility to gypsy moth defoliation. Forest Science 25:468–474.

Valentine, H.T., and D.R. Houston. 1984. Identifying mixed-oak stand susceptibility to gypsy moth defoliation: an update. Forest Science 30:270–271.

Walker, S.H., and D.B. Duncan. 1967. Estimation of the probability of an event as a function of several independent variables. Biometrika 54:167–179.

Science teachers of middle school and junior and senior high school students will be interested in a new publication of the Illinois Natural History Survey, *Legacy of a Pest: An STS Curriculum Guide for Understanding and Dealing with Biological Problems*. Materials in this guide cover a wide range of topics, including general entomology, forest ecology, and the impact and management of an introduced insect—the gypsy moth.

Manuscripts of high quality dealing with any aspect of natural history will be considered for publication in one of the Illinois Natural History Survey series: *Bulletin*, *Biological Notes*, *Circular*, and *Special Publication*. Authors who are not employees of the Survey are required to pay printing costs. Manuscripts should follow the recommendations of the third edition of the *Council of Biology Editors Style Manual* except that journal names in literature cited are to be spelled in full. The Survey expects to publish only one or two manuscripts by non-Survey authors yearly.

Send three copies of manuscripts to be considered for publication to the Office of the Chief, Illinois Natural History Survey, Natural Resources Building, 607 East Peabody Drive, Champaign, Illinois 61820. Before a manuscript is accepted for publication in the *Bulletin* or in *Biological Notes*, two or more outside referees must recommend it.

Illinois Natural History Survey

A Division of the Illinois Department of Energy and Natural Resources